z/OS Performance "HOT" Topics

Kathy Walsh

IBM

March 2, 2015

Session Number: 16964

#SHAREorg

SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.

Copyright (c) 2014 by SHARE Inc. Except where otherwise noted, this work is licensed under

Agenda

- Performance and Capacity Planning Topics
  - Introduction of z Systems z13 Processor
    - SMT
    - SIMD
  - Capacity Planning
  - CPU MF and HIS Support
  - zPCR Latest Status
  - zEDC Compression
  - New Performance White Papers
  - New RSM and WLM APARS
  - New SMF 42 Support
  - Interrupt Delay Time

- Addendum
  - Older APARs or Performance Information

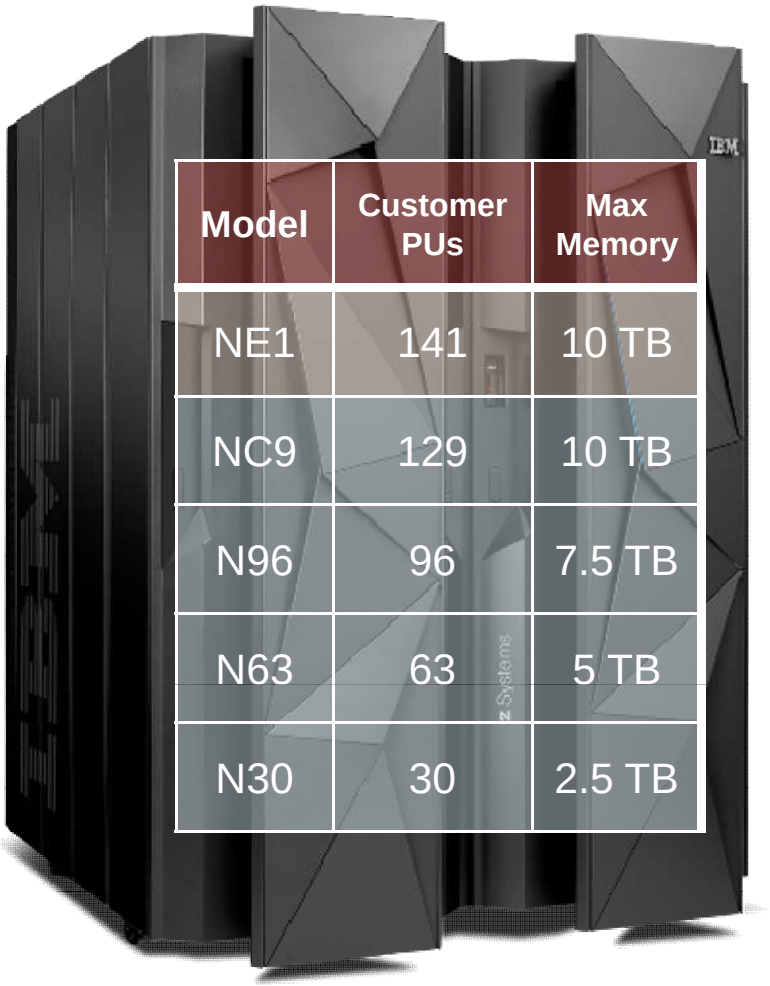
© 2015 IBM Corporation

www.ibm.com/support/techdocs

© 2015 IBM Corporation

z13 Overview

z13



Model	Customer PUs	Max Memory
NE1	141	10 TB
NC9	129	10 TB
N96	96	7.5 TB
N63	63	5 TB
N30	30	2.5 TB

- Machine Type
  - 2964
- 5 Models
  - N30, N63, N96, NC9 and NE1
- Processor Units (PUs)
  - 39 (42 for NE1) PU cores per CPC drawer
  - Up to 24 SAPs per system, standard
  - 2 spares designated per system
  - Dependant on the H/W model - up to 30, 63, 96, 129,141 PU cores available for characterization
    - 85 LPARs, increased from 60
  - Sub-capacity available for up to 30 CPs
    - 3 sub-capacity points
- Memory
  - RAIM Memory design
  - System Minimum of 64 GB
  - Up to 2.5 TB GB per drawer
  - Up to 10 TB for System and up to 10 TB per LPAR (OS dependant)
    - LPAR support of the full memory enabled
    - 96 GB Fixed HSA, standard
    - 32/64/96/128/256/512 GB increments
  - Flash Express
- I/O
  - 6 GBps I/O Interconnects – carry forward only
  - Up to 40 PCIe Gen3 Fanouts @ 16 GBps each and Integrated Coupling Adapters @ 2 x 8 GBps per System
  - 6 Logical Channel Subsystems (LCSSs)
    - 4 Sub-channel sets per LCSS
- Server Time Protocol (STP)

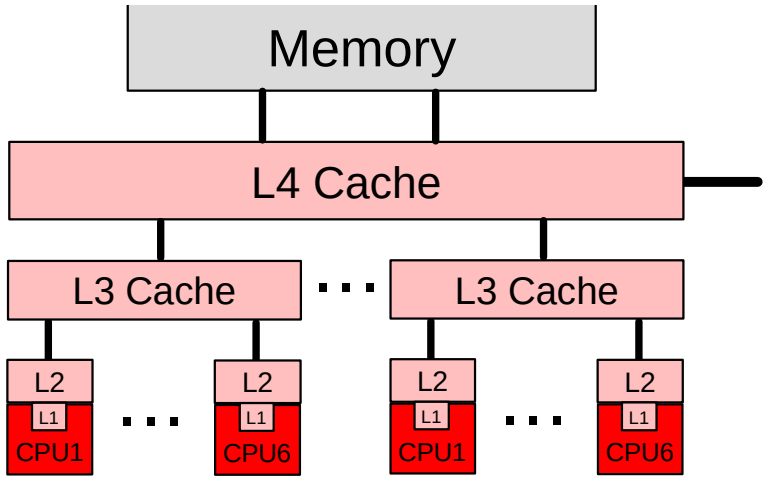
© 2015 IBM Corporation

IBM z13 versus zEC12 Hardware Comparison

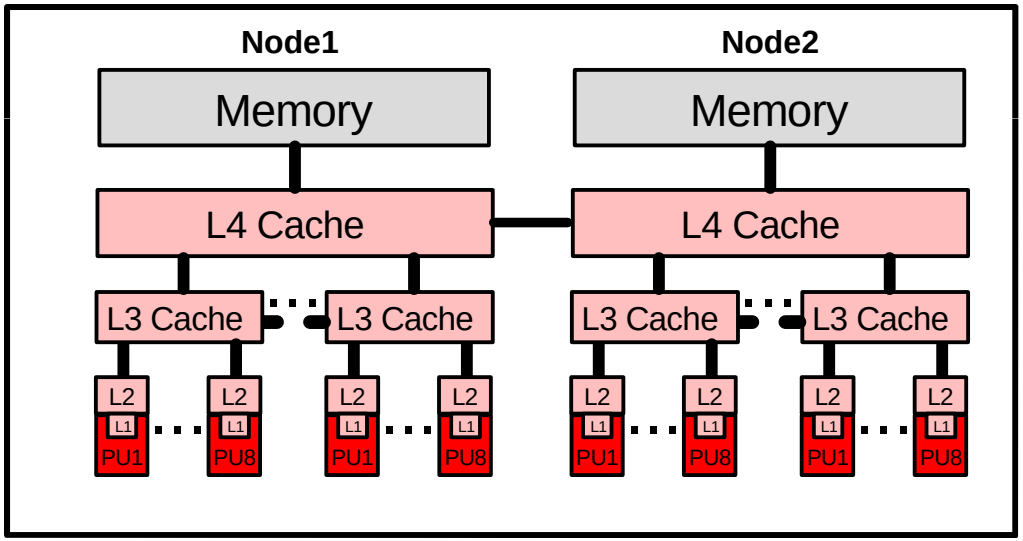
- zEC12
  - CPU
    - 5.5 GHz
    - Enhanced Out-Of-Order
  - Caches
    - L1 private 64k i, 96k d
    - L2 private 1 MB i + 1 MB d
    - L3 shared 48 MB / chip
    - L4 shared 384 MB / book
- z13
  - CPU
    - 5.0 GHz
    - Major pipeline enhancements
  - Caches
    - L1 private 96k i, 128k d
    - L2 private 2 MB i + 2 MB d
    - L3 shared 64 MB / chip
    - L4 shared 480 MB / node

      - plus 224 MB L3 NIC Directory

Single Book View



Single Drawer View - Two Nodes



© 2015 IBM Corporation

z13 Processor and Memory Assignment and Optimization

- Default processor assignments by POR, MES adds, and On Demand activation:
  - Assign IFLs and ICFs to cores on chips in “high” drawers working down, CPs and zIIP in low drawers working up
  - Objective: Keep partitions using IFLs and ICFs “away” from z/OS partitions on CPs and zIIPs in different drawers if possible

- PR/SM assigns available memory and logical processors at activation
  - Logical Processor specified in Image Profile assigned a *core* if Dedicated or if shared a *“home” drawer, node and chip*.
    - If Logical Processor becomes a HiperDispatch “Vertical High”, the Shared Logical Processor is assigned a specific core
  - Ideally assign all memory in one drawer with the processors if everything “fits”, with memory striped across drawers with processors if memory or processors must be split

© 2015 IBM Corporation

What is Simultaneous Multithreading (SMT)?

- Today each z System CPU supports a single instruction stream
  - z System workloads tend to receive a nontrivial # of cache misses
  - CPU generally unproductive while resolving cache miss
- z13 SMT makes CPU (Core) productive during cache misses
  - z13 supports two way SMT (Two instruction streams [Threads]) per core
    - Each Thread has its own unique state information (Registers, PSW, etc.)
    - Cannot necessarily execute instructions instantly and must compete and win the use of desired core resources shared between threads
    - z13 insures that one thread can't lock out the other
  - z13 allows following engine types to run in SMT mode
    - zIIPs under z/OS
    - IFLs under z/VM
- READY TO RUN Threads share core
  - Threads NOT READY TO RUN still unproductive while resolving cache miss
  - Core resources are productive when either READY TO RUN Thread is executing

© 2015 IBM Corporation

Simultaneous Multithreading (SMT)

- Simultaneous multithreading allows instructions from one or two threads to execute on a zIIP or IFL processor core.
- SMT helps to address memory latency, resulting in an overall capacity* (throughput) improvement per core
- Capacity improvement is variable depending on workload. For AVERAGE workloads the estimated capacity* of a z13:
  - zIIP is 38% greater than a zEC12 zIIP
  - IFL is 32% greater than a zEC12 IFL
- SMT exploitation: z/VM V6.3 + PTFs for IFLs and z/OS V2.1 + PTFs
- Notes:
  1. SMT is designed to deliver better overall capacity (throughput) for many workloads. Thread performance (instruction execution rate for an individual thread) may be faster running in single thread mode.
  2. Because SMT is not available for CPs, LSPR ratings do not include it

*Which approach is designed for the highest volume** of traffic? Which road is faster?*

***Two lanes at 50 carry 25% more volume if traffic density per lane is equal*

*Capacity and performance ratios are based on measurements and projections using standard IBM benchmarks in a controlled environment. Actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload .

© 2015 IBM Corporation

Exploiting SMT on z13 with zIIPs

- z/OS V2R1 SMT APARs must be applied
  - OA43366 (BCP), OA43622 (WLM), OA44439 (XCF)
  - z/OS manages threads according to the SMT Mode
- Define a LOADxx PROCessor VIEW (PROCVIEW) CORE|CPU for the life of the IPL
  - PROCVIEW CORE in z/OS on z13 enables SMT support
  - Fall back from PROCVIEW CORE to PROCVIEW CPU requires IPL
- New IEAOPTxx parameter to control zIIP SMT mode
  - MT_ZIIP_MODE=1|2
    - MT_ZIIP_MODE=2 for 2 active threads
    - When PROCVIEW CPU is specified the processor class MT Mode is always 1
  - Update IEAOPTxx to change mode and issue SET OPT=xx command
  - Requires HD=YES
- New LOADxx and IEAOPTxx controls ONLY available on z/OS V2R1 and higher

© 2015 IBM Corporation

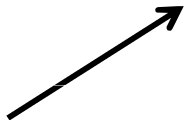
RMF CPU Activity – zIIP CPU Activity

```
                                                C P U  A C T I V I T Y

            z/OS V2R1                    SYSTEM ID SYSD            DATE 02/25/2015
                                         RPT VERSION V2R1 RMF      TIME 10.45.00
CPU         2964   CPC CAPACITY  4452          SEQUENCE CODE 000000000008DA97
MODEL       736    CHANGE REASON=NONE          HIPERDISPATCH=YES
H/W MODEL   N96
---CPU---    ---------------- TIME % ----------------   --- MT % ----   LOG PROC
NUM  TYPE    ONLINE    LPAR BUSY    MVS BUSY   PARKED   PROD   UTIL     SHARE %

 A    IIP    100.00    88.46         75.43       0.00    89.45  79.13    100.0  HIGH
                                      70.30       0.00
 B    IIP    100.00    81.96         68.84       0.00    87.79  71.95    100.0  HIGH
                                      59.34       0.00
 C    IIP    100.00    59.18         49.67       0.00    88.09  52.13     53.2  MED
                                      41.59       0.00
 D    IIP    100.00    31.78         27.10       0.00    88.09  27.99     53.2  MED
                                      22.31       0.00
 E    IIP    100.00    11.13          9.68       0.00    91.11  10.14      0.0  LOW
                                       8.33       0.00
TOTAL/AVERAGE           54.50         43.26               88.91  48.27    306.4
------------ MULTI-THREADING ANALYSIS ---------------
 CPU TYPE    MODE      MAX CF            CF       AVG TD
      CP        1       1.000         1.000        1.000
      IIP       2       1.384         1.225        1.585
```

© 2015 IBM Corporation

RMF CPU Activity – Workload Activity

```
                                W O R K L O A D   A C T I V I T Y                                    PAGE    1
      z/OS V2R1              SYSPLEX WSCZPLEX          DATE 02/25/2015          INTERVAL 04.59.999   MODE = GOAL
                             RPT VERSION V2R1 RMF      TIME 10.45.00

                                    POLICY ACTIVATION DATE/TIME 02/24/2015 17.37.33

------------------------------------------------------------------------------------------------------------------- SERVICE POLICY

REPORT BY: POLICY=WLM2013
                  WSC WLM 2013 policy

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---  --PROMOTED--  ----STORAGE----
AVG       96.73 ACTUAL                22  SSCHRT 202.7  IOC     800326  CPU 1012.133  CP       7.89  BLK     0.000  AVG     25289.43
MPL       96.73 EXECUTION             22  RESP     3.0  CPU      75322K  SRB     2.206  AAPCP    0.00  ENQ     0.000  TOTAL    1677310
ENDED     58207 QUEUED                 0  CONN     0.5  MSO          0  RCT     0.005  IIPCP    2.88  CRM     0.000  SHARED   7812.45
END/S    194.02 R/S AFFIN              0  DISC     0.0  SRB     164186  IIT     0.341                LCK     1.684
#SWAPS     122 INELIGIBLE              0  Q+PEND   0.7  TOT      76286K  HST     0.000  AAP      N/A  SUP     0.000  -PAGE-IN RATES-
EXCTD        0 CONVERSION              0  IOSQ     1.8  /SEC    254288  AAP       N/A  IIP   238.72                SINGLE       0.0
AVG ENC  30.40 STD DEV                15                                IIP   991.030                                BLOCK        0.0
REM ENC   0.00                                           ABSRPTN  2629                                SHARED       0.0
MS ENC    0.00                                           TRX SERV 2629                                HSP          0.0
```



Service Time is in Normalized SMT-1 time

IIP APPL % = 991.030 *100 / (299.999 * 1.384) = 238.69%

© 2015 IBM Corporation

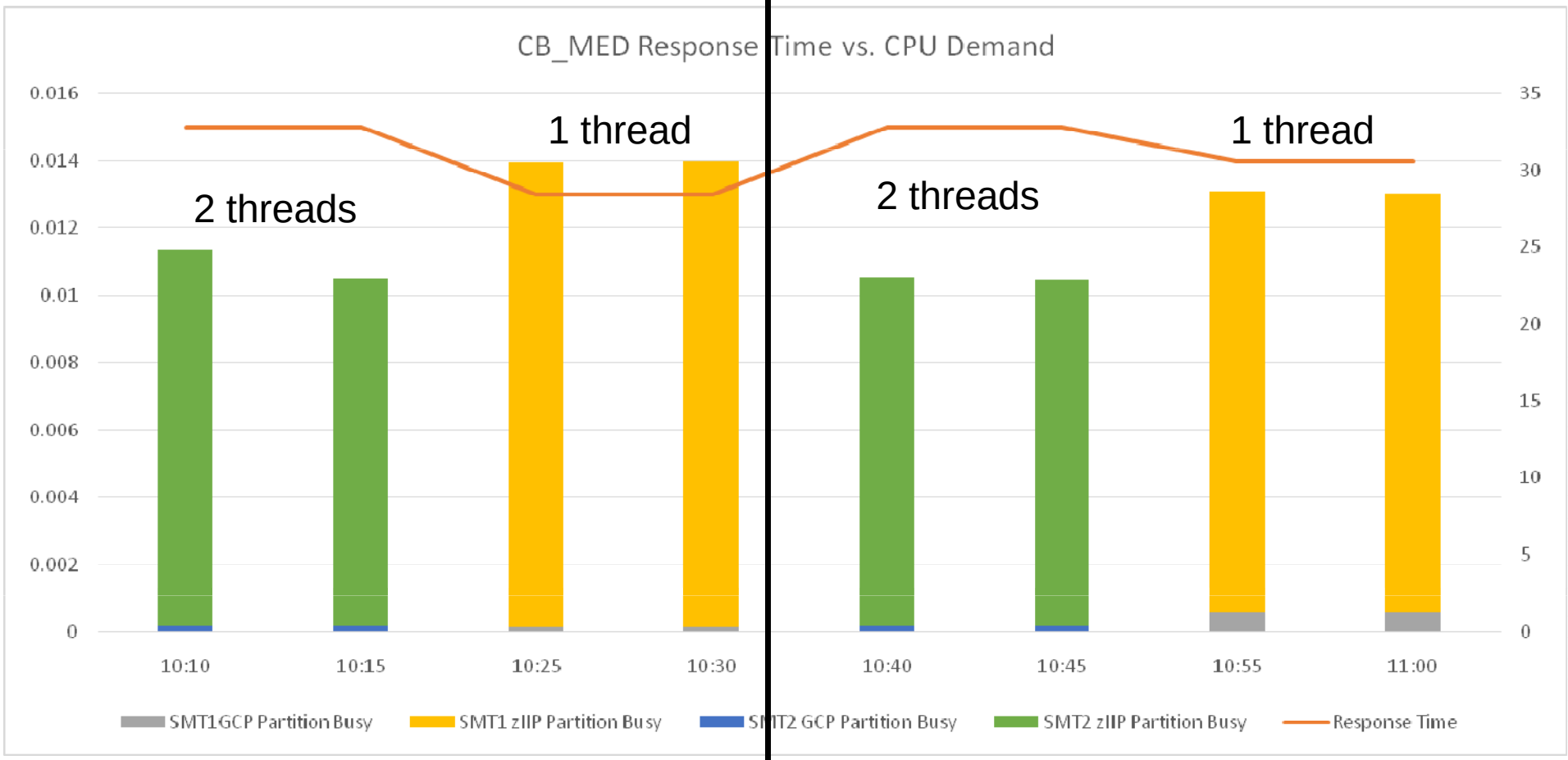
RMF CPU Activity – PDR Changes

```
                                P A R T I T I O N  D A T A  R E P O R T

    z/OS V2R1                 SYSTEM ID SYSD            START 02/25/2015-10.45.00  INTERVAL 000.04.59
                              RPT VERSION V2R1 RMF      END   02/25/2015-10.50.00  CYCLE 1.000 SECONDS

MVS PARTITION NAME                    UOSP02       NUMBER OF PHYSICAL PROCESSORS          96
IMAGE CAPACITY                           742                 CP                           36
NUMBER OF CONFIGURED PARTITIONS           85                 IFL                          48
WAIT COMPLETION                           NO                 IIP                          12
DISPATCH INTERVAL                    DYNAMIC

--------- PARTITION DATA ----------------  -- LOGICAL PARTITION PROCESSOR DATA --   -- AVERAGE PROCESSOR
                  -PROCESSOR  ----DISPATCH TIME DATA----   LOGICAL PROCESSORS  --- PHYSICAL PROCESSORS ---
NAME       S  WGT   NUM  TYPE   EFFECTIVE      TOTAL       EFFECTIVE    TOTAL  LPAR MGMT  EFFECTIVE  TOTAL
UOSP02     A  240     5  IIP   00.13.33.129  00.13.37.513       54.21    54.50      0.12      22.59  22.71
UOSP01     A  700    11  IIP   00.00.01.307  00.00.01.342        0.04     0.04      0.00       0.04   0.04
*PHYSICAL*                                   00.00.00.564                          0.02                 0.02
                             ------------  ------------                          -----      -----  -----
  TOTAL                      00.13.34.437  00.13.39.421                          0.14      22.62  22.76
```

Session 16816: RMF: The Latest and Greatest - Thur – 3:15 PM
Session 16930: z13 SMT (R)Evolution – Wed

© 2015 IBM Corporation

Early SMT Testing



HonorPriority=NO

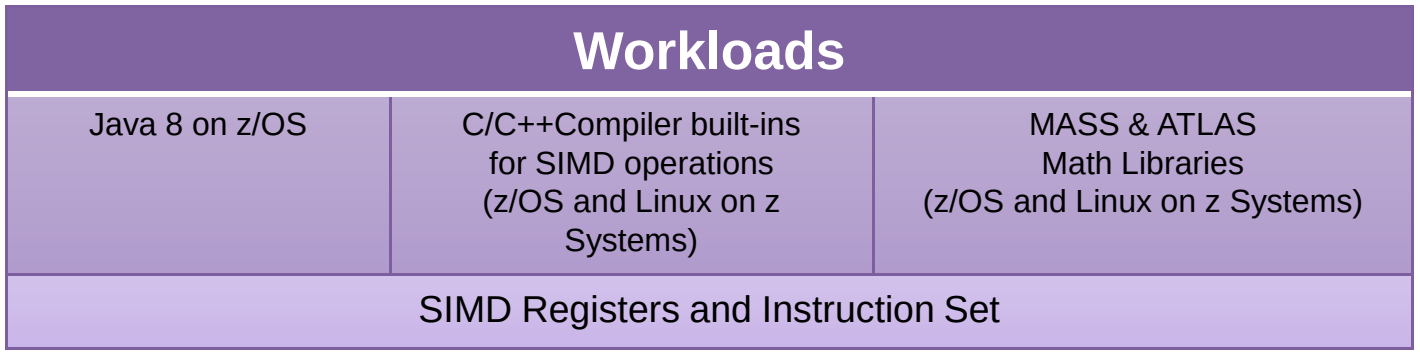
HonorPriority=YES

© 2015 IBM Corporation

Single Instruction Multiple Data (SIMD) Vector Processing

- Single Instruction Multiple Data (SIMD)
  - A type of data parallel computing that can accelerate code with integer, string, character, and floating point data types
- Providing optimized SIMD math & linear algebra libraries to minimize effort on the part of middleware/application developers
- Providing compiler built-in functions so software applications can leverage as needed (e.g. for use of string instructions)
- OS / Hypervisor Support:
  - z/OS: 2.1 SPE available at GA
  - Linux: IBM is working with its Linux Distribution partners to support new functions/features
  - No z/VM Support for SIMD
  - LOADxx MACHMIG can be used in z/OS to disable SIMD at IPL time
  - Compiler exploitation
    - IBM Java 8 => 1Q2015
    - XL C/C++ on zOS => 1Q2015
    - XL C/C++ on Linux on z => 2Q2015
    - Enterprise COBOL => 1Q2015
    - Enterprise PL/I => 1Q2015

MASS - Mathematical Acceleration Sub-System
ATLAS - Automatically Tuned Linear Algebra Software

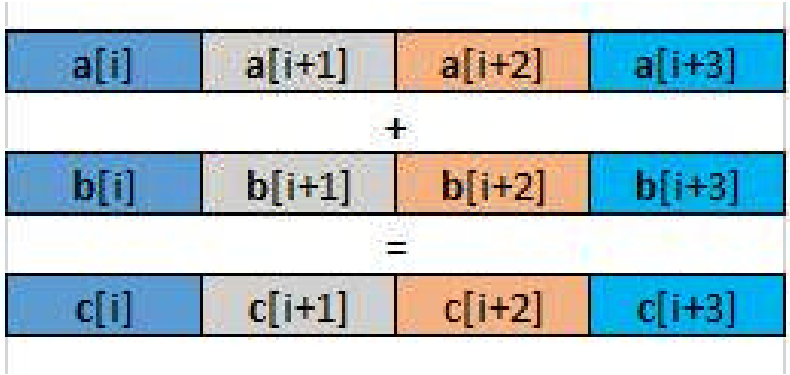


© 2015 IBM Corporation

IBM z13: SIMD – Single Instruction Multiple Data

Hardware for exploiting data-parallelism

- Large uniform data-set that needs the same operation performed on each element
- Can offer dramatic speedup to data-parallel operations (matrix ops, string processing, etc)



4x less iterations in the SIMD loop

```
//SISD C example, adding two arrays
for (i=0;i<128;i++)
    {
    c[i] = a[i] + b[i];
    }
```

```
//SIMD C example, adding two arrays
for (i=0;i<32;i++)
    {
    vec_add(c[i*4], a[i*4], b[i*4]);
    }
```

© 2015 IBM Corporation

IBM z13: SIMD – Single Instruction Multiple Data

IBM z13 running Java 8 on z/OS
Single Instruction Multiple Data (SIMD) vector engine exploitation

java.lang.String exploitation
- compareTo
- compareToIgnoreCase
- contains
- contentEquals
- equals
- indexOf
- lastIndexOf
- regionMatches
- toLowerCase
- toUpperCase
- getBytes

java.util.Arrays
- equals (primitive types)

String encoding converters
For ISO8859-1, ASCII, UTF8, and UTF16
- encode (char2byte)
- decode (byte2har)

Auto-SIMD
- Simple loops
  (eg. Matrix multiplication)

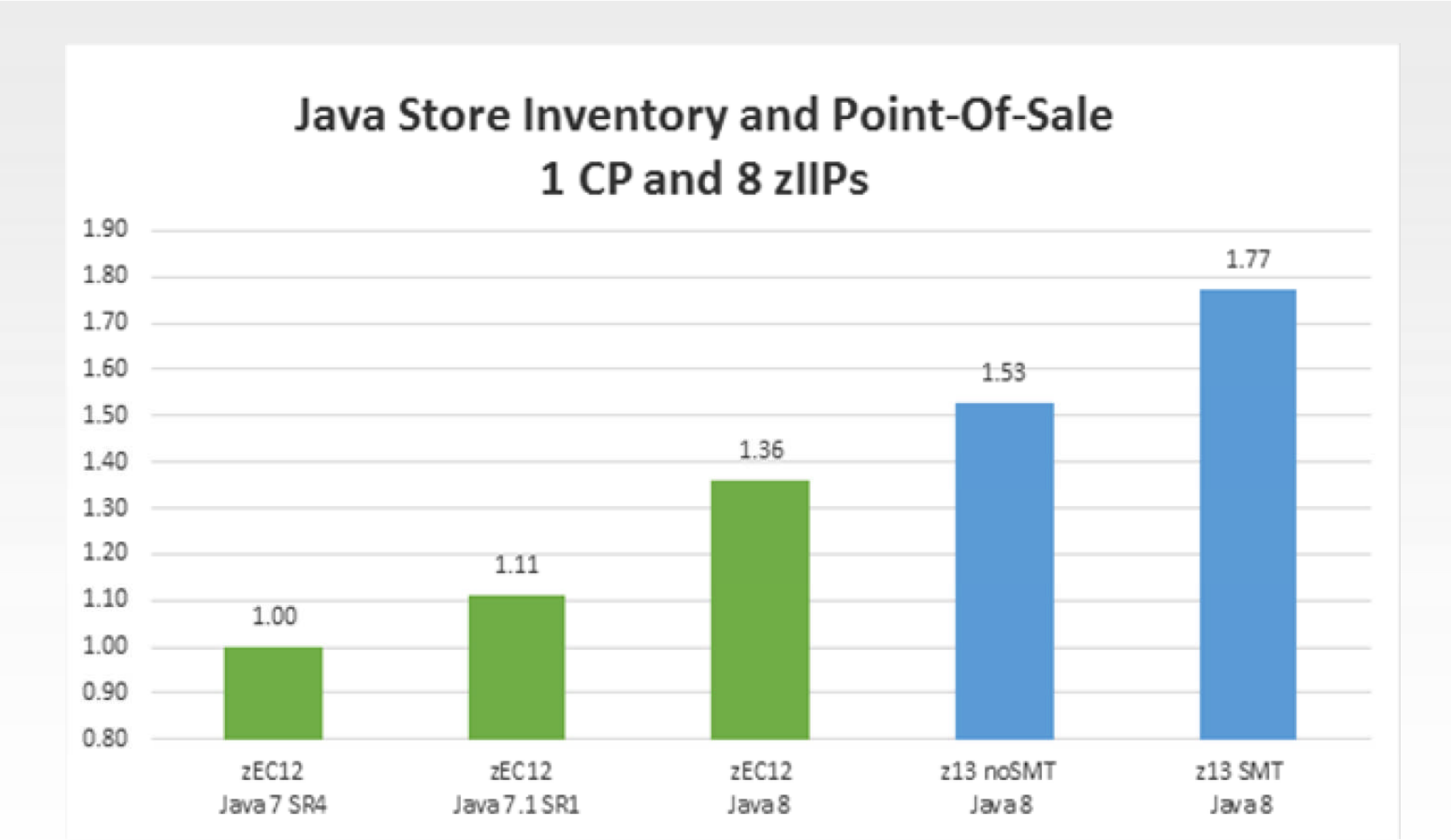
Primitive operations are between 1.6x and 60x faster with SIMD

(Controlled measurement environment, results may vary)

© 2015 IBM Corporation

Java on IBM z13

Java Store Inventory and Point-of-Sale App with IBM Java 8 and z13

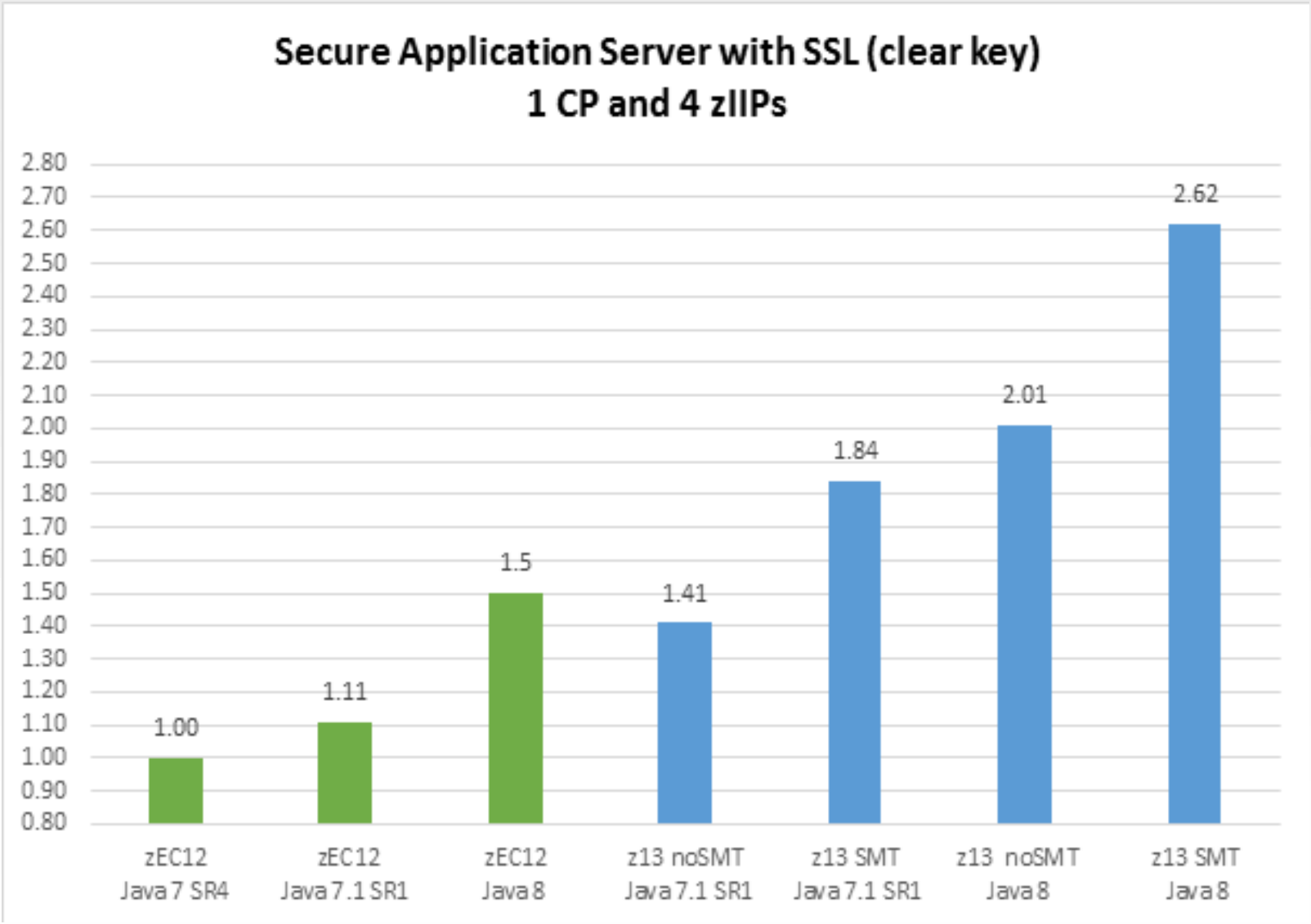


1.77x improvement in throughput with IBM Java 8 and IBM z13

(Controlled measurement environment, results may vary)

© 2015 IBM Corporation

Application Serving – SSL-Enabled DayTrader3.0



2.62x improvement in throughput with IBM Java 8 and IBM z13

(Controlled measurement environment, results may vary)

© 2015 IBM Corporation

Selecting a Workload for Capacity Planning

- New Processor Design
  - Includes major pipeline enhancements and Larger Caches
  - 1.10x (10%) average performance improvement at equal Nway Vs zEC12
- Workload Variability
  - Workloads moving to z13 may see more variability than the last few migrations
  - Potential Sources of Variability
    - Workload interaction with Processor Design – may have variable but unpredictable benefit to Instructions Per Cycle (IPC)
    - PR/SM placement of CPs and memory for an LPAR
- Critical to correctly specify the LSPR workload for an LPAR
  - Need to use CPU MF Counter data
    - Generates SMF 113 records with RNI and L1MP information
    - Captured by partition
  - CP3KEXTR utility will process SMF 113 data into an EDF file which can be used by zPCR (one file per partition)
  - If SMF 113 data is not available the LSPR Average workload category is used

Session 16814: The Relatively New LSPR and the IBM z13 - Tue – 11:15

© 2015 IBM Corporation

CPU Measurement Facility

- Available on all System z processors since the z10
- Facility provides hardware instrumentation data for production systems
- Two Major components
  - Counters
    - Cache and memory hierarchy information
    - SCPs supported include z/OS and z/VM
  - Sampling
    - Instruction time-in-CSECT
- z/OS HIS started task
  - Gathered on an LPAR basis
  - Writes SMF 113 records
- z/VM Monitor Records
  - Gathered on an LPAR basis – all guests are aggregated
  - Writes new Domain 5 (Processor) Record 13 (CPU MF Counters) records
- Minimal overhead

Session 16803: 2015 CPU MF Update - Tue – 10:00

© 2015 IBM Corporation

zPCR Latest Status

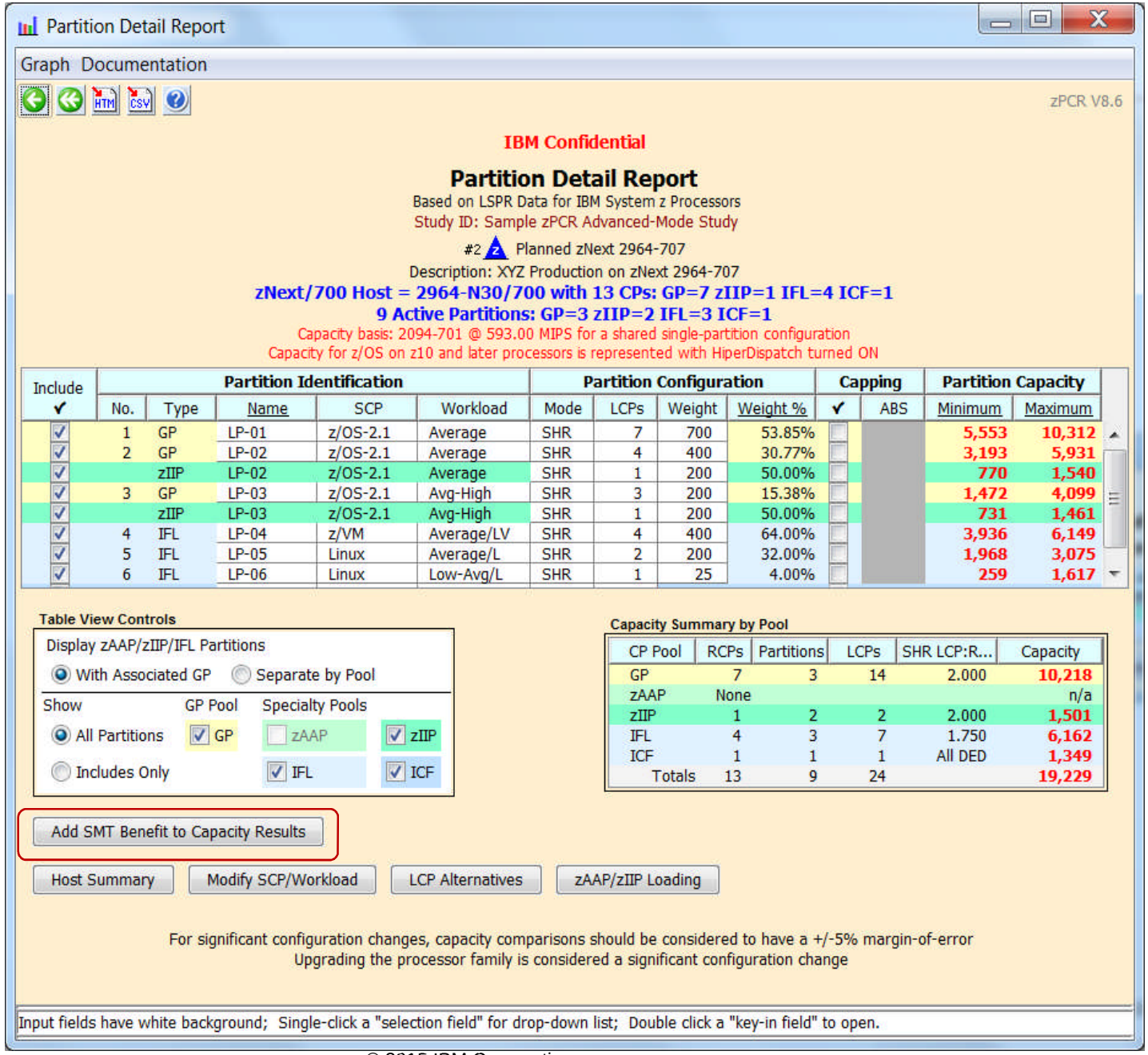
- Why use zPCR
  1. **LSPR Processor Capacity Ratios Tables**
  2. **LPAR Configuration Capacity Planning**

- Version 8.7c (2/16/2015)
  - The IBM z Systems (z13) processor family has been added, with 231 General Purpose models (90 sub-capacity and 141 full-speed) and 141 IFL models
  - LSPR data is now based on z/OS 2.1
  - New support added for Absolute Capping
  - On z13 processors SMT can be activated for z/OS running on zIIPS and z/VM running on IFLs. SMT is currently supported only by z/OS 2.1 and z/VM 6.3
  - In Advanced-Mode, the number of definable LPAR configurations has increased from 7 to 10

Session 16806: zPCR Capacity Sizing Lab – Part 1 Intro and Overview - Tue – 10:00 AM
Session 16798: zPCR Capacity Sizing Lab - Part 2 Hands-on Lab – Tue – 6:00 PM

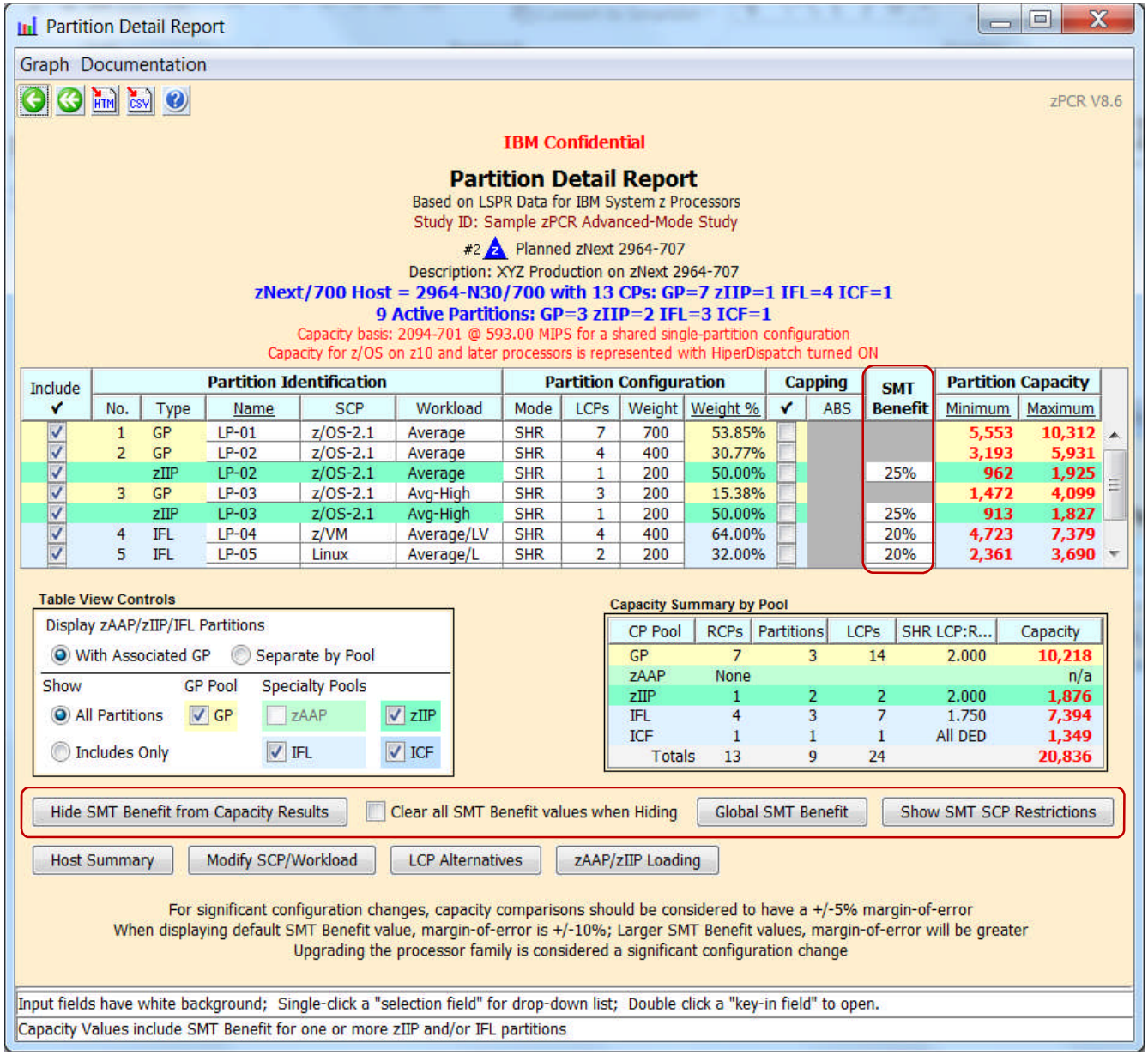
© 2015 IBM Corporation

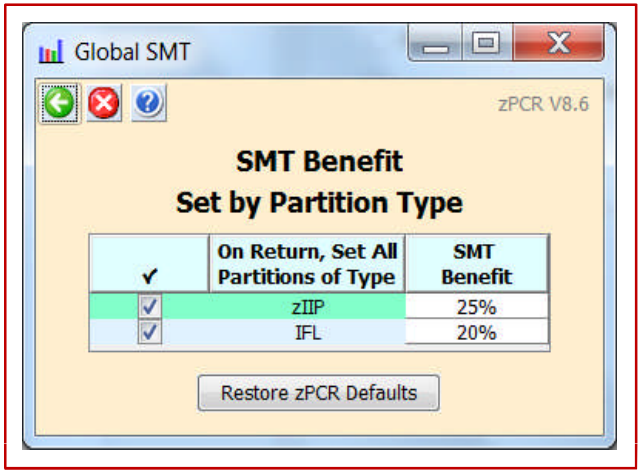
SMT Support in zPCR



© 2015 IBM Corporation

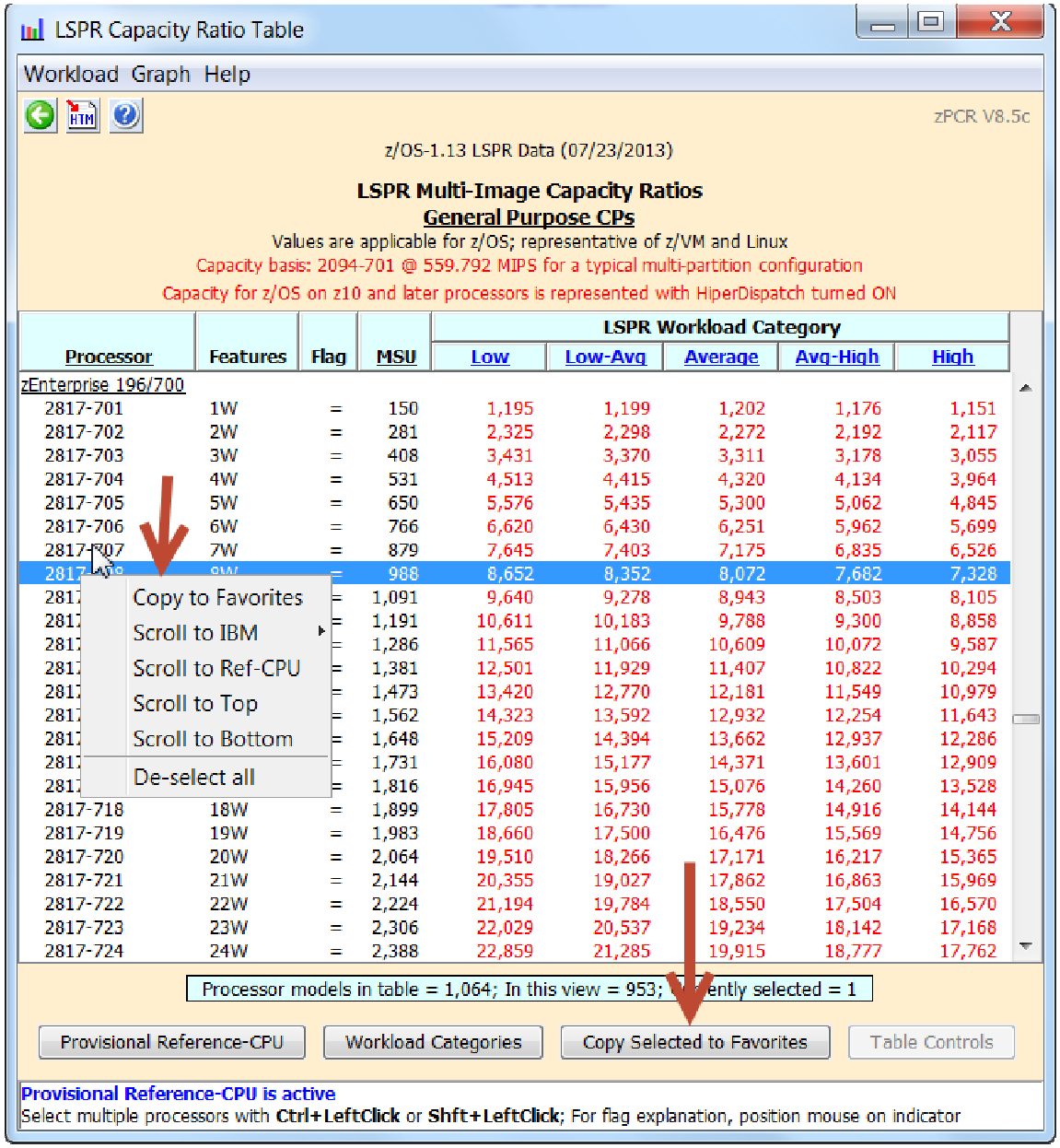
SMT Support in zPCR

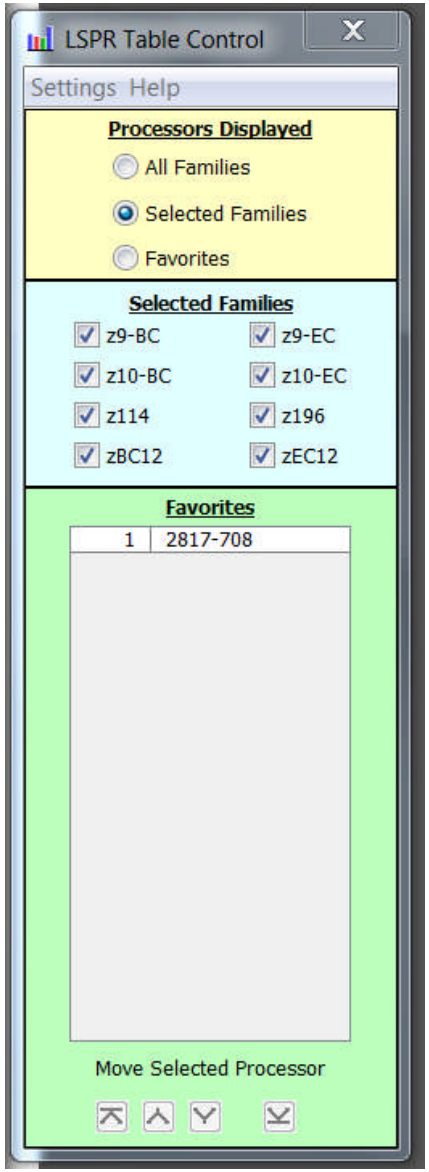




© 2015 IBM Corporation

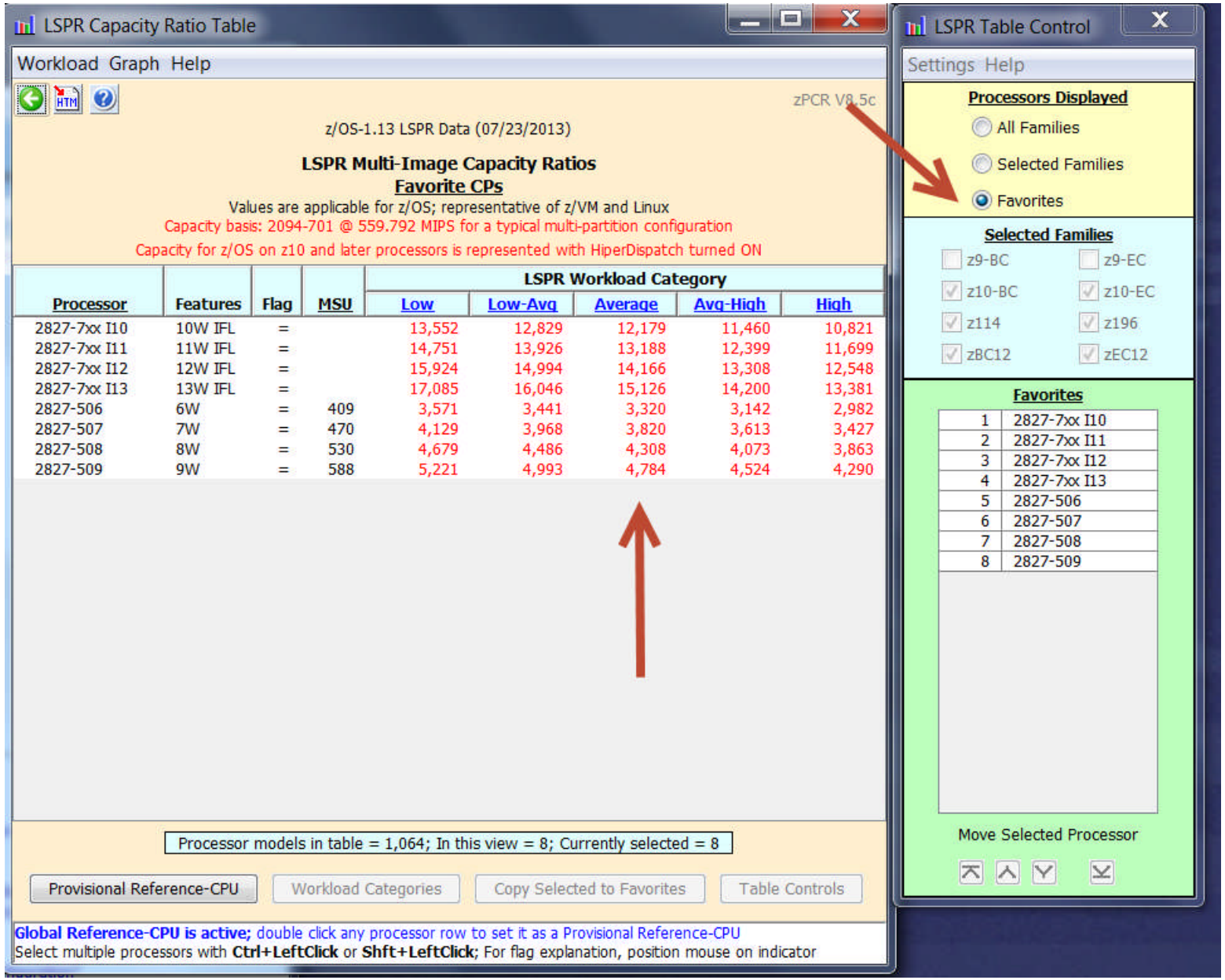
Example of Using zPCR Favorites Support





© 2015 IBM Corporation

zPCR Favorites Support



© 2015 IBM Corporation

Benefits of Exploiting Large Memory for DB2 Buffer Pools

- Three papers will be in the series, two available now:
  1. Advantages of Configuring More Memory for DB2 Buffer Pools V2
     - **WP102464**
     - http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102464
  2. Performance Report on Exploiting Large Memory for DB2 Buffer Pools with SAP®
     - **WP102461**
     - http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102461
  3. Benefits of very large buffers from a DB2 perspective (target date: 1Q2015)

- Concept:
  - DB2 buffer pools and associated memory were scaled at generally high (>80%) CPU utilization. Buffer pools were grown by as little as 14GB to 100's of GB
  - Performance benefits are highly workload dependent, many clients should see measureable, single digit percentage reduced CPU time and double digit percentage better response time
  - Large CPU time reductions due mainly to avoiding hardware cache disruptions, which are directly tied to the task switching associated with I/O operations. Thus, eliminating synchronous database IO through configuration of larger buffer pools can have a direct influence on CPU time used

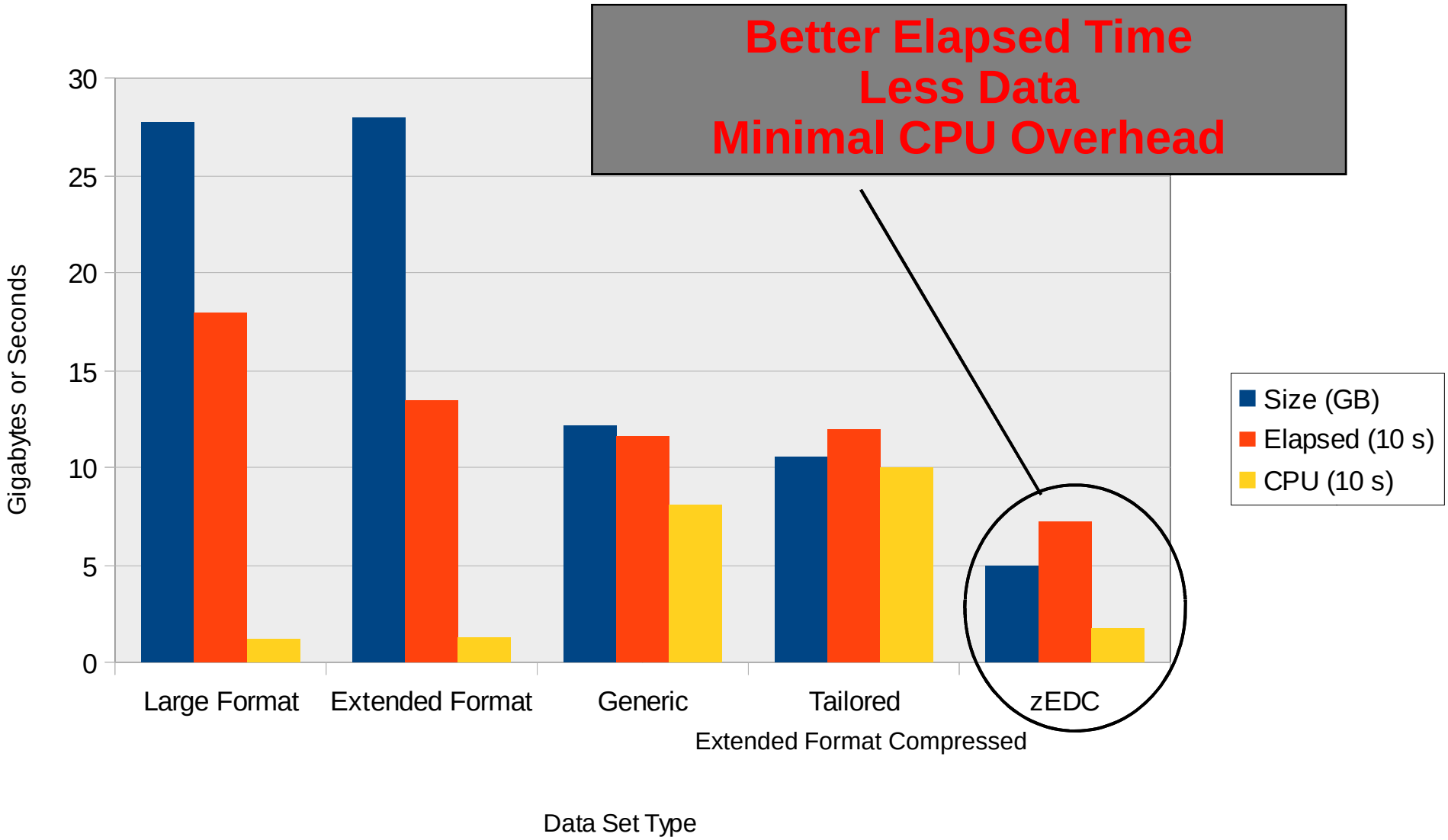
© 2015 IBM Corporation

zEC12 - Data Compression Express - zEDC

- The z Enterprise Data Compression (zEDC) Express offering provides a low-cost data compression to z/OS system services and applications
  - Implemented as a Peripheral Component Interconnect Express (PCIe) device
  - The compression function is provided via FPGA firmware
  - Can install up to 8 devices in a single processor
  - Each device is sharable by up to 15 LPARs
- Exploiters will see the following benefits
  - Increased performance for reading and writing compressed da
  - Reduced disk space
- Exploiters:
  - SMF Logstreams
  - Extended Format BSAM and QSAM data sets
  - IBM SDK for z/OS, Java Technology Edition, V7 R1
  - IBM Sterling Connect:Direct for z/OS Standard Edition V5.2
  - IBM Encryption Facility for z/OS
  - z/VM 6.3 support for guest exploitation of zEDC

© 2015 IBM Corporation

QSAM/BSAM zEDC – Value!



Disclaimer: Based on projections and/or measurements completed in a controlled environment.
Results may vary by customer based on individual workload, configuration and software levels.

© 2015 IBM Corporation

RMF Support for zEDC Express

- New support in RMF will provide information on PCI Express based functions
  - zEnterprise Data Compression (zEDC) capability using zEDC Express
  - RDMA (Remote Direct Memory Access) over Converged Enhanced Ethernet
- Information added via new subtype to SMF 74 (74.9) called PCI Express Based Function Activity
- Written by RMF Monitor III
  - General PCIE Activity – both zEDC and RDMA activity
  - Hardware Accelerator Activity
  - Hardware Accelerator Compression Activity
  - Types of data provided:
    - I/O Queue Time
    - I/O Execution Time
    - Number of compressed bytes in and out
    - Number of decompressed bytes in an out
    - Device drive buffer statistics

© 2015 IBM Corporation

RSM APAR - OA44207 – New Function

- Relieve scalability issues for LPARs running with large amounts of real storage between 256GB to 1 TB, or with a large LFAREA
  - Reduces amount of time RSM runs disabled during RSM initialization
  - Real storage used to contain the PFTE for the LFAREA is moved to a new area to help prevent depletion of preferred storage below the bar
  - Size of PLAREA is reduced when the LFAREA is very large allowing more preferred storage above the bar
  - New structures are defined for managing LFAREA frames which will reduce amount of time RSM runs disabled
  - Improvements to the Config STOR, offline command
  - RSM frame steal is improved by skipping over areas not eligible for steal

RSM APAR - OA44436 – New Function

- Support for LPARs running with large amounts of real storage
  - Automatic reconfigurable storage reduction when an RSU parm may cause insufficient preferred frames for IPL. MSG IAR0004I issued.
  - Indicator bit in RCE to determine when all online real frames at IPL have been initialized and available for use
  - Support for OA45505 to allow SRM to gather more granular global steal stats

© 2015 IBM Corporation

New RSM APAR - OA41968

- New IEASYSxx LFAREA parameter INCLUDE1MAFC
  - LFAREA=(64M,INCLUDE1MAFC)
  - Specifies the 1 MB pages are to be included in the available frame count (RCEAFC)
- RSM changes to:
  - Performs less paging when there is an abundance of available fixed 1M pages
  - More often break up fixed 1M pages to satisfy 4K page demand
  - Attempt to coalesce broken up fixed 1M pages when there is fixed 1M page demand, no guarantee coalescing will be successful
- RMF APAR in Support - OA42510
  - RMF PTFs must be applied prior to specifying INCLUDE1MAFC
  - RMF uses the RCEAFC to generate some of their reports and not applying OA42510 may lead to incorrect RMF reports
- Application programs:
  - Can check the RCEINCLUDE1MAFC bit to determine if the installation specified INCLUDE1MAFCin their LFAREA specification
  - When using STGTEST SYSEVENT to get information about the amount of storage available in the system if INCLUDE1MAFC is specified, available fixed 1M pages are included in this amount
- In a future IBM z/OS release, fixed 1M pages will be unconditionally included in the available frame count regardless of whether the INCLUDE1MAFC value is specified or not

© 2015 IBM Corporation

New WLM APAR - OA44526

- BLWLINTHD enhancements
- New support for blocked workloads
  - Allows lower threshold to be set
  - Defaults remain the same
- Useful for all online environments, with little to no batch workload, and heavy use of DB2
  - Helps prevent CPU starved workloads from holding locks which impact higher priority work
- Use RMF Workload Activity Report to measure the amount of blocked workload activity

```
W O R K L O A D   A C T I V I T Y

          --PROMOTED--
       BLK      3.240
       ENQ      0.000
       CRM      0.000
       LCK      0.000
       SUP      0.000
```

Check and understand why there are CPU times in any service classes

Session 16818: WLM Update for z/OS V2.2 and V2.1- Wed – 4:30 PM

© 2015 IBM Corporation

New SMF 42 Subtype 5 and 6 Support

- OA44322 and OA44319
- Updates for SMF records
  - SMF 42 subtype 5 - Storage Class / VTOC / VVDS I/O
  - SMF 42 subtype 6 - DASD Data Set Level I/O Statistics
- New Metrics Provided
  - Read and Write counts for High Performance FICON (zHPF)
  - Improved Resolution for Average Times reported
    - 1 microsecond units vs 128 microsecond units
  - Command Response Delay and Device Busy Delay
    - Subsets of PEND time

Note: These statements represent the current intention of IBM. IBM reserves the right to change or alter the plans in the future. IBM development plans are subject to change or withdrawal without further notice. Any reliance on this statement of direction is at the relying party's sole risk and does not create any liability or obligation for IBM.

© 2015 IBM Corporation

New SMF 42 Subtype 6 Support - WSC Example

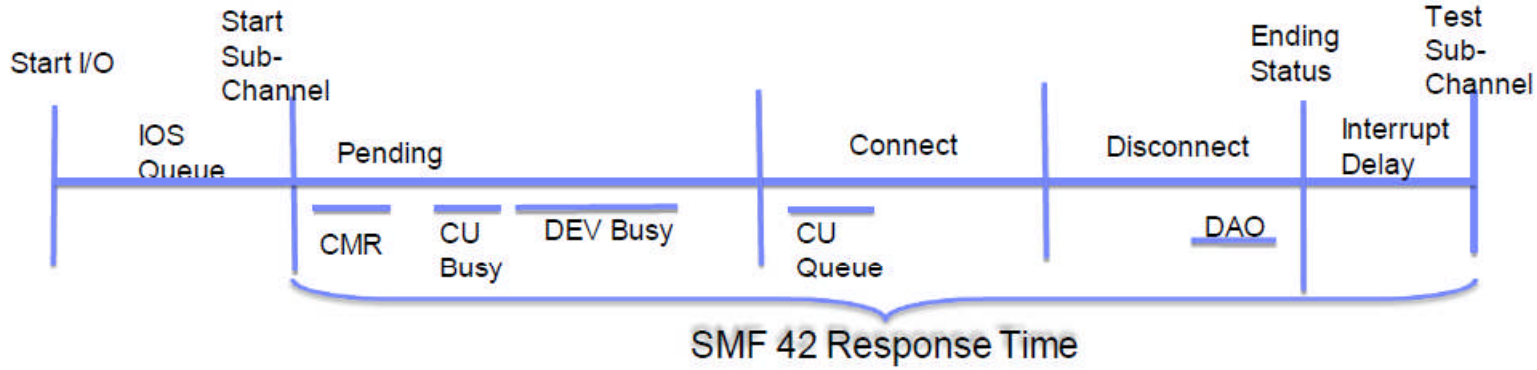
					"New" zHPF Metrics	
			Number of SSCHs	Number of Reads	**zHPF Reads**	**zHPF Writes**
DSN	**VOLSER**	**ADDR**	**ION**	**RDT**	**HRD**	**HWR**
WSCID34.ISPF.ISPPROF	ZOS048	9E05	1	1	0	0
WALSH.CPSTOOLS.JCL	ZOS050	9E06	1	1	0	0
SYS1.PARMLIB	OSPPGE	9C03	2	2	0	0
SYS1.BRODCAST	OSPPGE	9C03	1	1	0	0
MCCOX.EAVTST.DATASET	EAVOL1	8000	1	1	0	0
MATERIA.OMPRSYSA.ENV	ZOS048	9E05	2	2	0	0
JPBURG.SCM.RMFOUT	ZOS049	9E06	10	10	0	0

"New" 1 Microsecond Unit Times				"Old" 128 Microsecond Unit Times				
Total Response	Connect	Pend	Disc	Total Response	Connect	Pend	Disc	**"New" / "Old" Total Resp**
R1U	**C1U**	**P1U**	**D1U**	**Resp**	**Con**	**Pend**	**Disc**	**Ratio**
MS	**MS**	**MS**	**MS**	**MS**	**MS**	**MS**	**MS**	
10.69	10.44	0.23	0.00	10.62	10.37	0.13	0.00	1.01
11.48	11.24	0.22	0.00	11.39	11.14	0.13	0.00	1.01
0.27	0.10	0.14	0.00	0.13	0.00	0.00	0.00	2.07
0.23	0.07	0.13	0.00	0.13	0.00	0.13	0.00	1.78
0.47	0.28	0.14	0.00	0.38	0.26	0.13	0.00	1.21
6.55	0.19	0.15	6.19	6.40	0.13	0.13	6.14	1.02
0.33	0.14	0.15	0.00	0.26	0.00	0.13	0.00	1.30

Note: These statements represent the current intention of IBM. IBM reserves the right to change or alter the plans in the future. IBM development plans are subject to change or withdrawal without further notice. Any reliance on this statement of direction is at the relying party's sole risk and does not create any liability or obligation for IBM.

© 2015 IBM Corporation

SMF 42 Control Unit & Channel Measurement Fields



- *Response time is elapsed time from Start Subchannel to TSCH; CPU dispatching can play a role*
- *Service time = Pend + Connect + Disconnect*
- *Pend time includes channel selection overhead (not explicitly reported), along with CMR delay, Device Busy delay, CU Busy delay, but not CU Queue delay*
- *Command Response delay: Starts when the channel sends the 1st command until it receives a response from the controller. One round trip through the fabric*
- *Device Busy delay measures time associated with an initial status of "device busy" delay time. Example: device reserved by another system*
- *CU Queue delay is time spent queued in the Controller. example: extent conflicts*
- *Device Active Only - Channel End to Device End on last command when not issued at the same time. Subset of Device-Defer Time. example: writes to Synchronous PPRC Secondary*

© 2015 IBM Corporation

LPAR I/O Interrupt Delay Time

- APAR OA37160 – New Function
  - z/OS 1.12 and above and zEC12, zBC12, or z13 processor
- New Interrupt Delay Time
  - Keeps track of the time from when a subchannel is made status pending with primary status to when the status is cleared by TSCH
  - Tracking the accumulated delay encountered due to PR/SM needing to dispatch z/OS to process the interrupt as well as any z/OS delay
- APAR OA39993 - RMF support
  - RMF 74.1 record (device) and RMF 79.9

© 2015 IBM Corporation

LPAR I/O Interrupt Delay Time

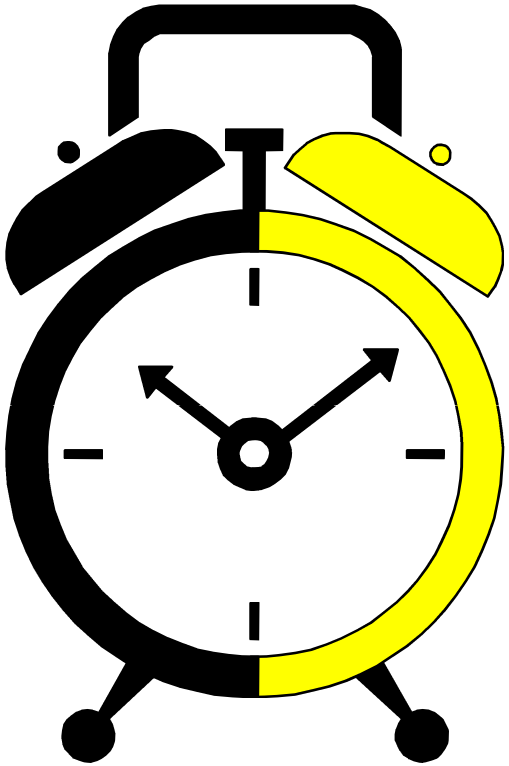
- Definition of response time in RMF is not impacted by this new field
- Data may be subject to a Low Utilization effect where large numbers are possible if I/O rates are very low
- Problem found in the channel microcode which causes incorrect values
  - MCL will be provided for the zEC12 and zBC12 by 2Q2015
- If I/O rates are > 1 or 2 per second and values are large (>0.2ms) then look at:
  - LPAR weight
  - CPENABLE
  - Processor Busy
  - z/OS Busy

STORAGE GROUP	DEV NUM	DEVICE TYPE	DEVICE ACTIVITY RATE	AVG RESP TIME	AVG IOSQ TIME	AVG CMR DLY	AVG DB DLY	AVG INT DLY	AVG PEND TIME	AVG DISC TIME	AVG CONN TIME
WSCPDQDB	402E	33909	0.000	.896	.000	.256	.000	1.41	.384	.000	.512
WSCPDQDB	402F	33909	0.000	.512	.000	.256	.000	.000	.384	.000	.128
WSCQAQDB	4030	33909	0.000	.128	.000	.000	.000	.000	.128	.000	.000
WSCPPOOL	4031	33909	0.160	6.31	.000	.060	.000	.118	.158	6.00	.153
WSCPPOOL	4032	33909	1.462	3.67	.000	.061	.000	.090	.159	3.17	.345
WSCPDQDB	403B	33909	0.000	1.02	.000	.256	.000	1.79	.384	.000	.640
WSCPPOOL	403C	33909	1.569	1.72	.000	.059	.000	.074	.154	1.00	.565
WSCPPOOL	403D	33909	1.117	4.20	.000	.039	.000	.098	.138	3.93	.131
WSCPDB	403E	33909	0.004	.689	.000	.030	.000	.069	.138	.354	.197

© 2015 IBM Corporation

Addendum

- Older information which should still be understood, or make you go Hmmmm.
- APARs which are still causing issues, even though they are old.



© 2015 IBM Corporation

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*BladeCenter®, DB2®, e business(logo)®, DataPower®, ESCON, eServer, FICON, IBM®, IBM (logo)®, MVS, OS/390®, POWER6®, POWER6+, POWER7®, Power Architecture®, PowerVM®, S/390®, Sysplex Timer®, System p®, System p5, System x®, System z®, System z9®, System z10®, Tivoli®, WebSphere®, X-Architecture®, zEnterprise®, z9®, z10, z/Architecture®, z/OS®, z/VM®, z/VSE®, zSeries®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries, or both.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

© 2015 IBM Corporation

Notice Regarding Specialty Engines (e.g., zIIPs, zAAPs and IFLs):

Any information contained in this document regarding Specialty Engines ("SEs") and SE eligible workloads provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs). IBM authorizes customers to use IBM SEs only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the “Authorized Use Table for IBM Machines” provided at:
www.ibm.com/systems/support/machine_warranties/machine_code/aut.html (“AUT”).

No other workload processing is authorized for execution on an SE.

IBM offers SEs at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

© 2015 IBM Corporation

White Paper: Checklist for Establishing Capacity Groups

- **WP102437**
  - Describes the HMC procedure for setting up a Group Capacity profile and assigning LPARs to become members of the group
  - Gives step by step directions including all of the screen shots needed to complete the task including updating the Running System
  - Shows how to validate via RMF the Capacity Group definitions

White Paper: Managing zOS on a Uni-processor - V2

- **WP100925**
  - Discusses managing work on a uniprocessor system including the impact of looping or CPU-intensive tasks
  - Availability features such as blocked workload and ENQ promotion

© 2015 IBM Corporation

z13 Processor and Memory Assignment and Optimization

- PR/SM optimizes resource assignment when triggered
  - Triggers: Partition activation or deactivation or significant processor entitlement changes, dynamic memory increases or processor increases or decreases (e.g. by CBU) or MES change
  - Optimization
    - Examine partitions in priority order by size of “processor entitlement” (dedicated processor count or shared processor pool allocation by weight)
    - Changes Logical Processor “home” drawer/node/chip assignment
    - Moves Logical Processor to different chips, nodes, drawers (LPAR Dynamic PU Reassignment)
    - Relocates partition memory to active memory in a different drawer(s) using Dynamic Memory Relocation (DMR)
      - If available but inactive memory hardware is present (e.g. hardware driven by Flexible or Plan Ahead) in a drawer where more active memory would help: activate it, reassign active partition memory to it, and deactivate the source memory hardware, again using DRM
      - PR/SM can use all memory hardware but concurrently enables no more memory than the client has paid to use

© 2015 IBM Corporation

Mobile Workload Pricing for z/OS

- Benefits
  - Improves the cost of growth for mobile transactions processed in System z environments such as CICS, IMS, DB2, MQ and WAS
  - Mobile Workload Pricing (MWP) for z/OS enhances Sub-Capacity pricing
    - Mitigates the impact of Mobile on MLC charges where higher transaction volumes cause a spike in processor utilization
    - Normalizes the rate of transaction growth
  - No infrastructure changes required, no separate LPARs needed
    - It is an enhanced way of reporting sub-capacity MSUs
    - System runs as it always has, workload execution is not altered
- Hardware requirements
  - Available to all enterprises running a zEC12 or zBC12 server (actual mobile work may run on any zEnterprise machine including z196 and z114)
- MWP Announcement Letters – 2014-05-06

• AP14-0219	Asia Pacific	• JP14-0279	Japan
• A14-0429	Canada	• LP14-0279	Latin America
• ZP14-0280	Europe, Middle East, Africa	• 214-223	United States

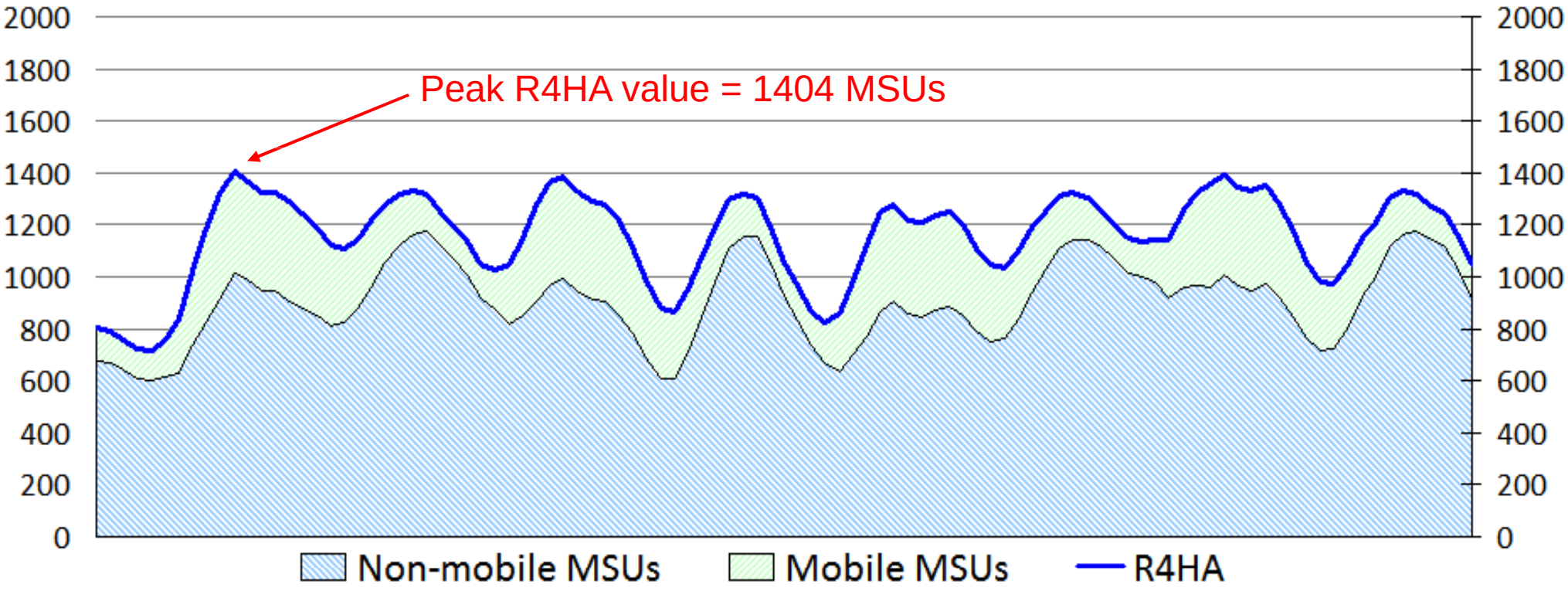
© 2015 IBM Corporation

Mobile Workload Pricing Reporting Process

- New Mobile Workload Reporting Tool (MWRT) – available 06/30/2014
  - A new Windows-based Java tool to report sub-capacity MSUs and make adjustments to reported LPAR MSUs based on Mobile transaction data
    - Standard SCRT methodology plus new feature to adjust for Mobile workload impact
    - New tool replaces SCRT for customers who take advantage of Mobile Workload Pricing
- Customers must track mobile transactions and produce a file showing mobile CPU consumption each month
    - Record mobile transaction data, including CPU seconds, on an hourly basis per LPAR
    - Load the resulting data file into MWRT each month (IBM-specified CSV format)
    - Run MWRT and submit the results to IBM each month (Replaces SCRT process)
- MSU adjustments and monthly peak calculation for billing
  - MWRT will **subtract 60%** of the reported Mobile MSUs from a given LPAR in each hour, adjusting the total LPAR MSU value for that hour
    - This will function like a partial “off-load” from a software pricing perspective
    - When LPAR value is adjusted, all software running in LPAR benefits from lower MSUs
    - Tool will calculate monthly MSU peak for a given machine using adjusted MSU values

© 2015 IBM Corporation

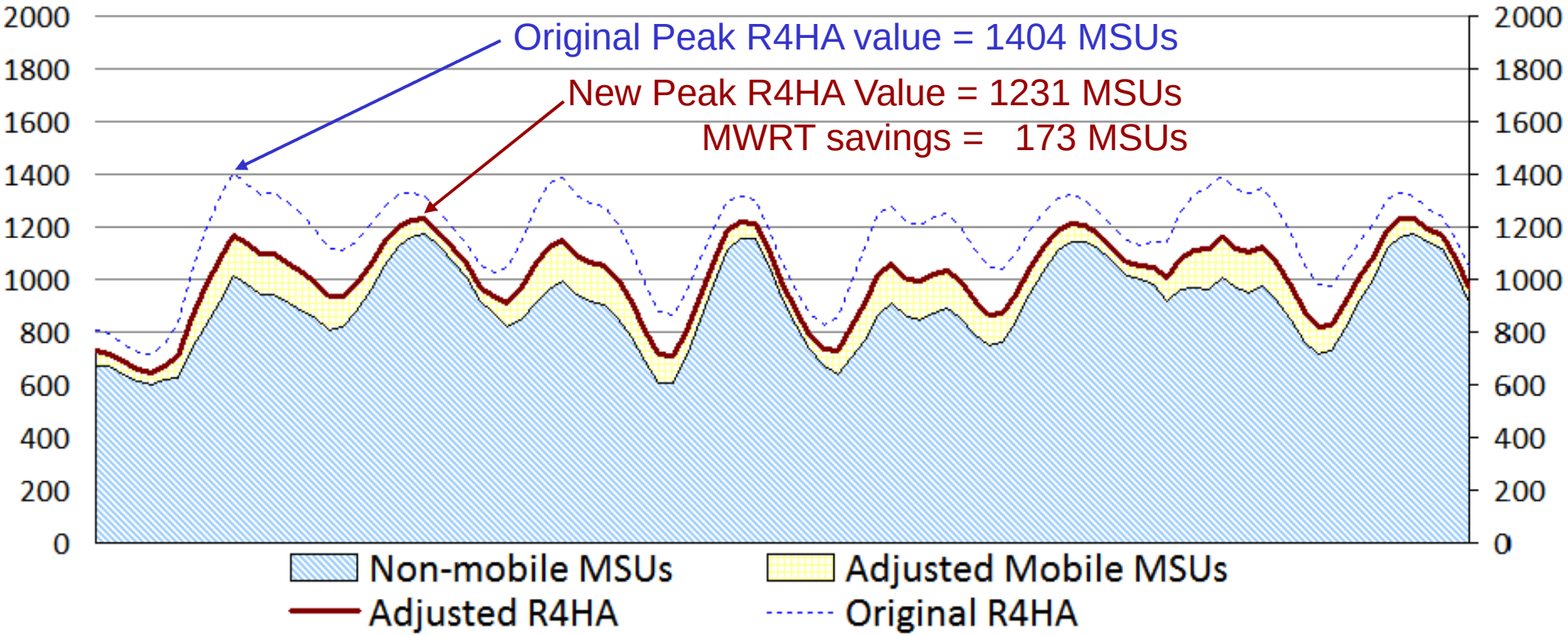
Example: Sample LPAR MSU values by hour



- SCRT calculates the Rolling 4-Hour Average (R4HA) MSU peak
  - All workloads are included

© 2015 IBM Corporation

LPAR MSU values adjusted for mobile contribution

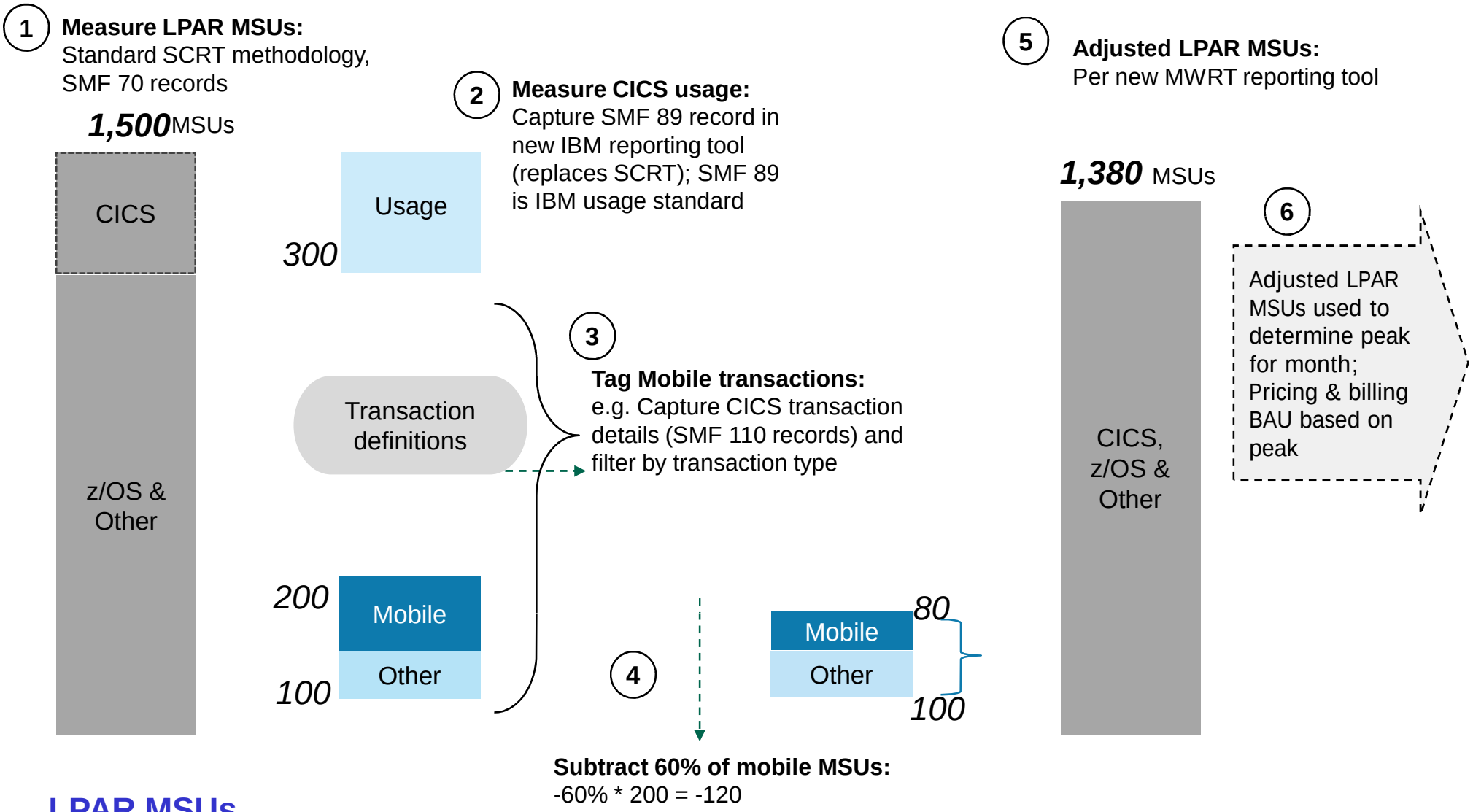


- MWRT removes 60% of the Mobile workload, interval-by-interval
  - Non-mobile workload is unchanged
  - Billing for the month is based upon the newly calculated R4HA curve after the mobile workload has been reduced

Provides benefit when Mobile workloads contribute to monthly peak MSUs; Off-peak MSU adjustments will not affect MSUs used for billing.

© 2015 IBM Corporation

Example: reducing Mobile impact to LPAR peak



LPAR MSUs for billing (Standard)

z/OS/Other 1,500

CICS 1,500

LPAR MSUs for billing (Adjusted)

z/OS/Other 1,380

CICS 1,380

Figures are for illustrative purposes only.

© 2015 IBM Corporation

Identifying Mobile Transaction Workload

- Customers are responsible for processing their mobile transaction data into a predefined format to be loaded into MWRT for each sub-capacity reporting period.
  - IBM must approve the data gathering methodology.
- The data must consist of **general purpose processor CPU seconds** for each mobile transaction program summarized by hour by LPAR for all machines processing mobile transactions.
  - Detailed instructions, including CSV file format, available in the MWRT user's guide.
- Mobile Workload Pricing Defining Programs:

5655-S97	CICS TS for z/OS V4	5635-A04	IMS V13
5655-Y04	CICS TS for z/OS V5	5655-DSQ	IMS DB VUE V12
5722-DFJ	CICS VUE V5	5655-DSM	IMS DB VUE V13
5635-DB2	DB2 V9 for z/OS	5655-L82	WS MQ for z/OS V6
5605-DB2	DB2 10 for z/OS	5655-R36	WS MQ for z/OS V7
5615-DB2	DB2 11 for z/OS	5655-W97	WS MQ for z/OS V8
5697-P12	DB2 VUE V9	5655-VUE	WS MQ VUE V7
5697-P31	DB2 10 VUE	5655-VU8	WS MQ VUE V8
5697-P43	DB2 11 VUE	5655-N02	WebSphere App Server for z/OS V7
5635-A02	IMS V11	5655-W65	WebSphere App Server for z/OS V8
5635-A03	IMS V12		

© 2015 IBM Corporation

Worksheet example

IBM MWP contract supplement Customer Worksheet

The information from this worksheet will be used to prepare the official contract Supplement (Z126-6628) to the IBM Mobile Workload Pricing Addendum (Z126-6300).

Mobile Application name:	Bank Account Mobile App
MWP Defining Program(s):	CICS, IMS

Mobile Application Details

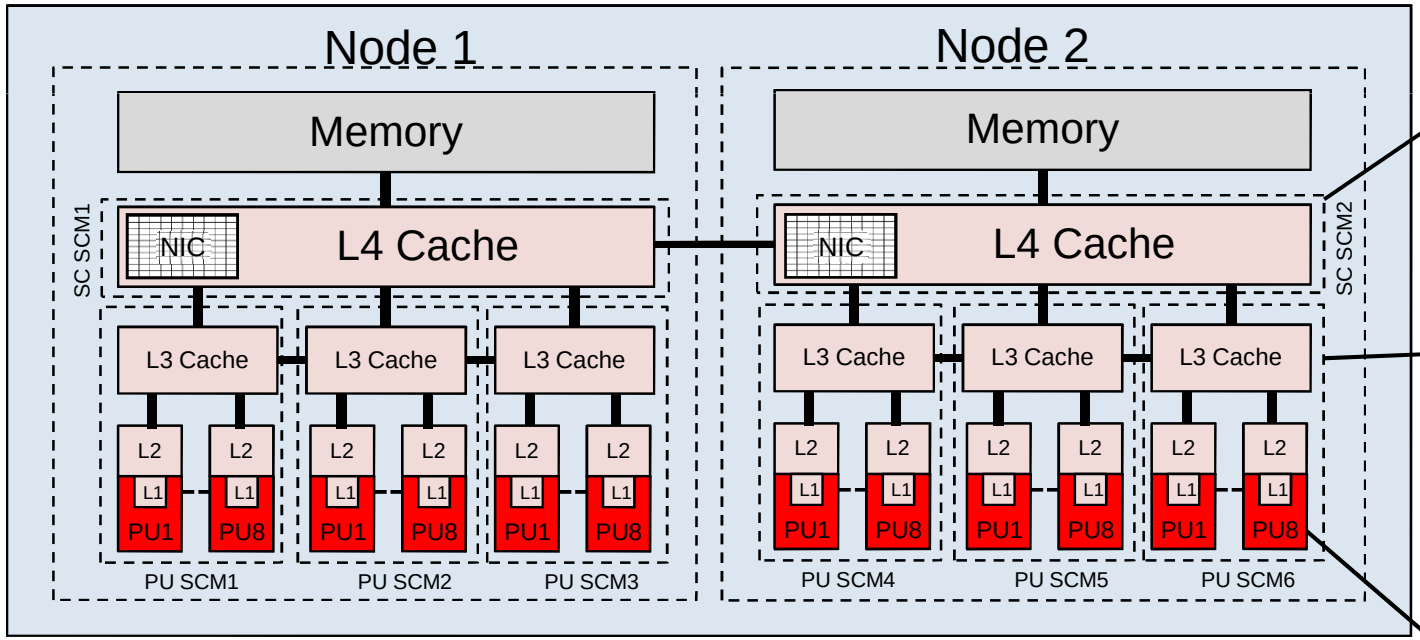
Details regarding the Mobile Application specified above are in the table below. Fill out one row of information in the table for each MWP Defining Program used by the Mobile Application named above.

A. MWP Defining Program	B. Data source	C. Method to distinguish mobile transactions from other workload	D. Client process for capturing and processing the mobile data
e.g., CICS	*e.g., SMF 110*	*e.g., All transaction types originating from mobile devices (e.g. iPhone app traffic) have been assigned unique names and routed to a specific region.*	*e.g., Filter mobile transaction types by name from all transactions using a SAS program, and sum the general purpose processor CPU seconds by hour for the affected LPARs.*
e.g., IMS	*e.g., IMS Logs*	*e.g., Enable IMS Transaction Level Statistics to produce 56FA log records for the mobile transactions; All mobile transactions contain the word "mobile" in the LUName.*	*e.g., Extract the mobile transaction details from the IMS log records using IMS PA and sum the general purpose processor CPU seconds by hour for the affected LPARs.*

© 2015 IBM Corporation

z13 CPC Drawer Cache Hierarchy Detail

Single CPC Drawer View (N30 Model) – 2 Nodes



2 SC SCM Chips
(1 per node)

6 PU SCM Chips
(3 per node)

* Up to 8 PU cores per chip

Single PU core

* Not all PU's active

Node 1 - Caches
- L1 private 96k i, 128k d
- L2 private 2 MB i + 2 MB d
- L3 shared 64 MB / chip
- L4 shared 480 MB / node
  - plus 224 MB NIC

Node 2 - Caches
- L1 private 96k i, 128k d
- L2 private 2 MB i + 2 MB d
- L3 shared 64 MB / chip
- L4 shared 480 MB / node
  - plus 224 MB NIC

© 2015 IBM Corporation